

Spatial Data Correlation: An interactive 3D visualisation tool for correlating the motion capture data streams from different devices

Carina E.I. Westling
School of Media and Film
University of Sussex
c.e.i.westling@sussex.ac.uk

Leon Barker
School of Informatics
University of Sussex
l.barker@sussex.ac.uk

Aoife Healy
CSHER
Staffordshire University
a.healy@staffs.ac.uk

Robert Needham
CSHER
Staffordshire University
a.healy@staffs.ac.uk

Harry J. Witchel
Brighton & Sussex Medical School
BN1 3PX
h.witchel@bsms.ac.uk

Nachiappan Chockalingam
CSHER
Staffordshire University
n.chockalingam@staffs.ac.uk

We have developed an interactive, three-dimensional visualisation that plots individual motion capture data streams so they can be visually inspected concurrently to yield a unified, multimodal view of a complex problem. The primary objective in this study is to facilitate analysis of groups of putatively correlated time series with large numbers of data points. Our unique contribution is that we have developed a three-dimensional data visualisation tool with fly-through data capabilities, in order to distinguish false correlations from genuine correlations.

Motion capture. Engagement. Cognitive ergonomics. Affective computing. Postural micro-movements. Posture recognition. Unity3D. 3D visualisation. Three-dimensional data visualisation.

1. INTRODUCTION

We have developed a visualisation technique using the games engine Unity 3D for comparing two sources of time series data based on multiple time series; we were interested in comparing data streams from many different time series to determine whether the two sources of data were, in principle, correlated with each other. The visualisation technique we have developed may be applied to a range of potential correlation problems, but we were specifically interested in validating the usefulness of one source of motion capture data (accelerometry) compared to data gathered simultaneously by a gold standard technology.

The 'gold standard' for motion capture is the Vicon opto-electronic system, which is a rapid, multi-camera computational system for three-dimensional position tracking of passive, infrared-reflective markers (Barker *et al.* 2006; Villanueva *et al.* 2000). The putative technology to record head movements is a tri-axial accelerometer placed on the crown of the head. Accelerometry has a range of potential advantages over camera-based technologies, including cost, portability, and ease of set-up, but accelerometers do not measure position; it measures acceleration (i.e. force). Here we used the raw accelerometry signal (which

responds clearly to gravity) to estimate head tilt, rather than deriving head position or head angles computationally by integrating acceleration. Traditional comparisons look for correlations between simultaneous recording methods and tend to summarise data using statistical cross-correlations. These cross-correlation methods can, however, find notable correlations where there are none, especially:

- When the time series data have many thousands of data points
- When the time series data include long periods without change in movement or activity
- When the correlations incorporate lags; or
- When the correlations incorporate linear transformations.

Raw accelerometer data has a range of issues that can make its use difficult. Accelerometry has drift, so between each three-minute data stream we had to zero the data; without running extensive calibration protocols between each run, this means that the relationship between the zero points of accelerometry and Vicon were poorly defined. Furthermore, Vicon measures position, while accelerometry measures acceleration (i.e. the second derivative of position), so the correlation

between the changes in raw accelerometer signal and Vicon signal are unknown without extensive calibration. For this reason we sought correlations between the changes in the two signals, and we used a best-fit linear transformation of the accelerometry signal in the absence of repeated calibrations every three minutes.

In this study, we compared the two technologies for recording head movement data of healthy volunteers whilst they watched or interacted with audio-visual stimuli in a seated position. Head position data of seated volunteers is a putative biomarker for vigilance and engagement during human-computer interaction (Balaban *et al.* 2004); head position while seated is characterised by long periods (~10 seconds) where the head remains still, interspersed with a range of small and slow movements that may be instrumental, emotional, fatigue-driven, or random.

We found excellent statistical correlations for the simultaneously gathered data, but we also found good statistical correlations when comparing data that were deliberately not matched; we used the reverse time series data as the negative control, because reverse data would have identical variability/stability to the experimental data. When comparing the Vicon data to the reverse time series negative control data, we found some good (but not excellent) correlations. Based on this, we decided to develop a data visualisation model that could facilitate distinguishing false correlations from real correlations by visual inspection. A three-dimensional grid has the advantage of allowing for coordination of several modalities, or comparison of larger datasets. By arranging multiple datasets as adjacent datapoint clouds in this interactive, three-dimensional grid, we can identify false correlations that otherwise can seem significant, based on statistical analysis alone.

2. METHODOLOGY

2.1 Measurement technologies

In this case study, we have compared motion capture data from naturalistic movements by a single volunteer, who was watching or interacting with a set of 12 three-minute audio-visual stimuli (films, photo montages, games and quizzes), presented on a computer monitor (Witchel *et al.* 2012). The volunteer was seated in a standard four-legged, chair with no armrests and cushioned fabric covering back and seat. The experiment (including interviews) lasted less than 90 minutes. A Vicon 8-camera opto-electric motion analysis system (OMG Plc., Oxford, UK) was compared to a head-mounted triaxial accelerometer (MIE Medical Research, Leeds, UK). The sampling rate for the

Vicon data and the accelerometer data was 100 Hertz.

We affixed spherical reflective markers and measurement devices to the participant, who was subsequently allowed to acclimatise to the devices for 15 minutes, while filling in forms and hearing the instructions. The participant wore a headband, allowing us to affix the accelerometer, and 4 reflective markers to mark head position. Other markers were placed on the shoulders, clavicle, and sternum.

2.2 Stimuli

We used 12 short stimuli (170-190 seconds), selected and/or designed to elicit responses ranging from highly disengaged, to very engaged. These stimuli included computer games, quizzes, short films, and photomontages. Each stimulus was preceded by 45 seconds of 'television snow' plus white noise, to establish a baseline signal before each stimulus, followed by a brief timing signal. The order of the stimuli was randomised. In previous experiments where we have used the same set of stimuli, we found that certain stimuli reliably received (mean) high VAS scores for interest, while others received high VAS scores for boredom, and others were ranked in between (Witchel *et al.* 2012).

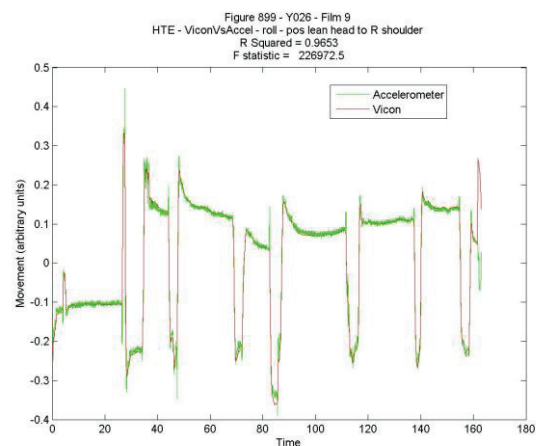


Figure 1: Vicon (red) and accelerometer (green) time series overlay, after linear transformation

2.3 Analysis

The two technologies that we used to make concurrent measurements were not recorded by the same device, so we assumed that, although we used timing signals throughout the experiment, there would be small discrepancies in time between different measurements. We did not attempt to calibrate all the measurements (e.g. the accelerometer) to the displacement measurements made by Vicon, as the two techniques measure different things. Instead we sought potential

correspondences by adjusting each signal to the Vicon signal by determining any delay, and then making a linear transformation of the signal in Matlab (MathWorks). The goodness of fit was represented by the R-squared term of the regression, but we used graphical inspection as the primary judgement for determining similarities in the signal pairs.

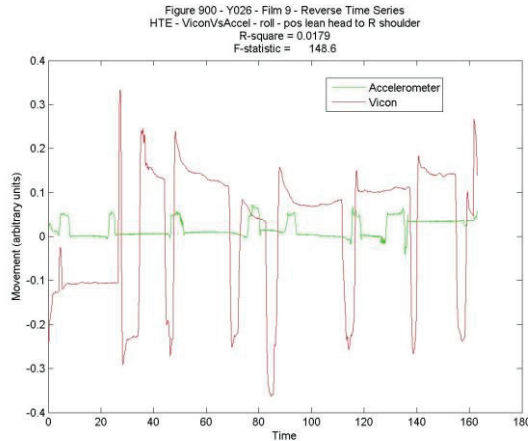


Figure 2: Vicon (red) and accelerometer (green) reverse time series (negative control signal pair) overlay. Note that the red signal is identical to figure 1, and that the green time series from figure 1 is reversed and inverted

With time series of 8000 points or more, and long stretches of steady readings interspersed with spikes of sudden change, unsynchronised signals with similar distributions (as judged by histograms) can appear statistically related. We compared each potentially synchronised signal pair to a negative control pair, with similar histograms but no true synchronisation.

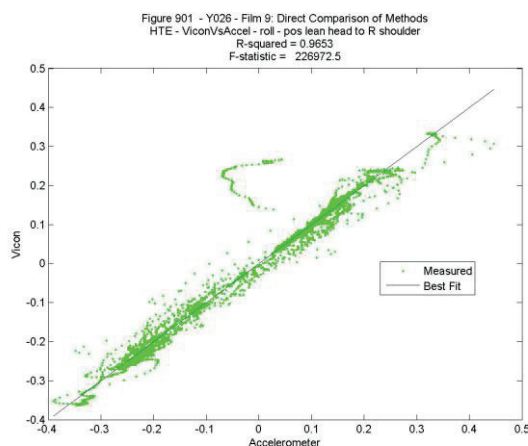


Figure 3: Correlation plot, showing excellent correlation between Vicon and accelerometer data

The negative controls were generated by reversing the non-Vicon (accelerometer) signal in its time course, and running this a priori unsynchronised signal pair through the regression and linear

transformations. The R^2 values for the negative controls never exceeded 0.4 (most were below 0.1), whereas the R^2 values for the statistically synchronised pairs were always greater than 0.9.

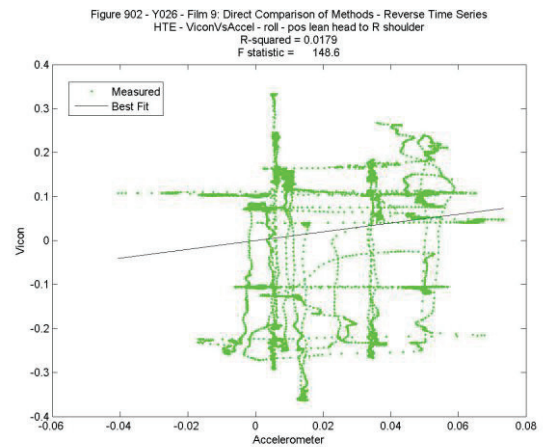


Figure 4: Correlation plot showing the box pattern of uncorrelated data from the negative control, generated by reversing the time series for the accelerometer signal

We initially graphed both the experimental and the negative control signal pairs (described in 2.3) in three ways: direct overlay of the two time series (figure 1 and 2 for the experimental and the negative control signal pair, respectively), direct comparison of the two signals at each time point (figure 3 and 4 for the experimental and the negative control signal pair, respectively), and a Bland Altman plot (figure 5 and 6 for the experimental and the negative control pair, respectively). The Bland Altman x-axis is the average between the two measurements for that time point; these are compared to how many of the differences between the two time series are 2 standard deviations away from the mean. Perfectly equivalent measurements generate a horizontal line at $Y=0$. (Bland and Altman, 2010).

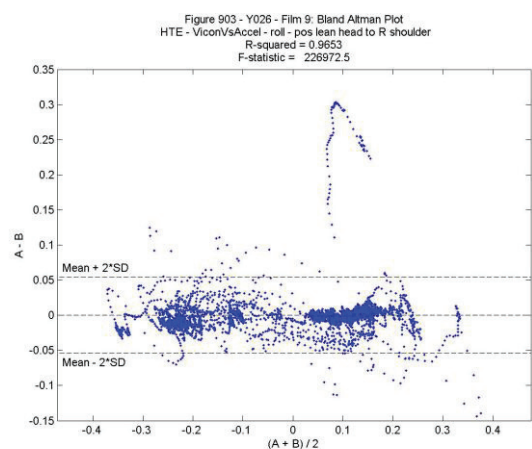


Figure 5: Bland Altman plot of experimental signal pair

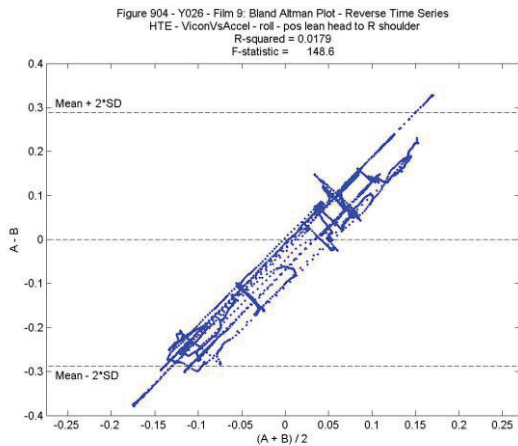


Figure 6: Bland Altman plot of negative control pair. This plot is identical in shape to figure 4 with its X axis foreshortened and rotated 45° clockwise

Graphing was automated in Matlab for each of the 12 stimuli for pitch, roll, and yaw.

2.4 3D Model

We used the games engine called Unity3D as the development environment for our data-visualisation. The engine facilitated the rendering of objects in three dimensions as well as the development of statistical algorithms using programming languages javascript and C sharp.

Through the development of a three-dimensional data visualisation tool we began to consider whether such visualisations present any benefit to statisticians during the qualitative analysis of large datasets. By creating an interactive model that allows data to be plotted and viewed from any angle we were able to compare and evaluate the quality of data captured earlier in our study.

The visualisation tool utilises primitive sphere within Unity3D to represent individual data-points. These data-points are rendered at runtime after the data is loaded from json arrays, which contain of the spatial coordinates. The tool created enables users to zoom into and move through the data cloud, each data-point can be inspected in close detail (figure 7).

Cameras can be located anywhere within the scene, and using either javascript or C sharp, camera functionalities can be extended so that the

user can interact with the model, or change viewpoints using mouse or cursor keys.

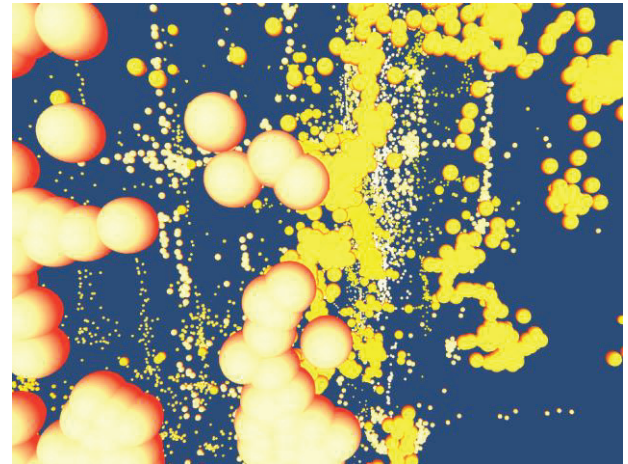


Figure 7: Camera functions allow you to zoom in and even move through the multiple correlated datapoint clouds, for closer inspection

3. RESULTS

After resampling the measurement pairs (the concurrently recorded Vicon and accelerometer signals), and identifying any delays, we performed a linear transformation based on a linear regression to make the signals similar in nature. We selected for this study the head pitch angle measurements, as they offered the best overlay fits. All the regression models had R^2 values greater than 0.92. We generated a corresponding negative control (see 2.3), and loaded the datasets into our 3D model, for visual inspection.

The graphical analysis of the accelerometer vs. Vicon signals was based on plotting for each time point the value of each measurement against the other. For this comparison, we used a correlation line plot with a slope of 1. As visible on the left in figure 8, showing the forward time series data plot with a slight counter-clockwise rotation, it has excellent correlation, whereas the reverse time series data (right in figure 8) shows the characteristic box pattern of uncorrelated data, even where it yields R^2 values of 0.3 to nearly 0.4. We found that three-dimensional visualisation of the datasets allowed us to correctly identify true and false correlations by visual inspection.

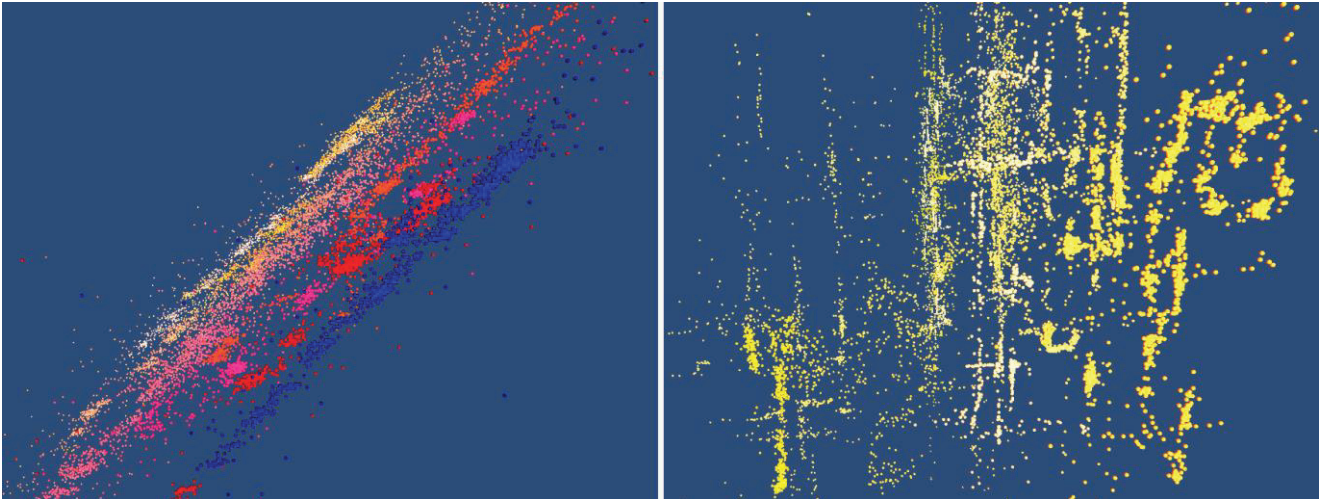


Figure 8: Shows correlated (gradient) and uncorrelated data (yellow) plotted in our 3D visualisation tool

4. DISCUSSION

Our presentation demonstrates how games engines can be utilised by statisticians to visually qualify and interrogate large and complex datasets. By creating a tool that enables data scientists to easily observe inconsistencies and correlations, our aim is to assist researchers in their pursuit of tangible outcomes. For the purpose of this example, we have used Unity3D to create a model that allows researchers to concurrently interact with and compare data from many time course experiments, in order to show that this can enhance and facilitate analysis of large and complex datasets.

While this demonstration shows how a three-dimensional comparison of multiple datasets, e.g. research data from interactions with a full stimuli set for one volunteer, or data from interactions with one stimulus for an entire cohort, can facilitate identification of true vs. false correlations, we are also interested in using our model for multimodal mapping. We are currently working toward three sets of composite motion capture modalities, in order to elucidate how the different modalities qualify each other. Our initial efforts to formulate a testable hypothesis for the interpretation of postural micro-movements as proxies for degrees of engagement suggest that the complexity of non-verbal signals needs a multimodal approach.

4.1 Postural measurements

Although postural micro-movements have been explored as a potential source of information on subconscious emotions and tendencies, many of the previous investigations pre-date computer-based motion capture. A range of investigators

(Calvo and D'Mello, 2010; D'Mello *et al.* 2011; D'Mello and Graesser, 2010; Mota and Picard, 2003) have more recently measured postural micro-movements of seated individuals, generally using chair-based pressure sensitive mats (Tekscan Inc., USA) to detect changes in centre of pressure. Head and neck movements, which are both more numerous and possibly more psychologically revealing, require other technologies that generally are head-mounted. Magnetic sensors (e.g. flock of birds, Ascension Technology) have been used for this purpose; these have the disadvantage of being sensitive to metallic distortion. Our own investigation of centre of force as a proxy for psychologically meaningful postural adjustments, by way of using force plates, have not so far yielded results that encourage coherent interpretation. We have recently initiated a comparative study of Vicon and Tekscan chair pressure mat signals, and we have not yet seen indications that the chair mat, in our hands, is capable of yielding measurements for postural adjustment that reveal head movements, or even hand gestures, possibly due to automatic, compensatory movements that balance the centre of force. We cannot, however, give any conclusive comparison until we have finished analysing our data.

The previous researchers, referred to above, have focused on continuous stimuli. Our research with discrete stimuli has uncovered the need for qualifying modalities when attempting to objectively measure engagement during interaction, as engagement manifests in different patterns according to the type of stimuli used, and the affordances offered by the stimuli and viewing regime. Our experiments are concerned with interaction with screen-based presentations, and in

many cases the affordances offered primarily invite the user to engage with focused attention, which is a state of relative stillness. This is particularly true with stimuli that do not offer affordances for physical interaction, but we also found it to be the case with many interactive stimuli. Further research is needed before we can make any conclusive claims about the distinctions between intellectual, emotional and physical forms of engagement, and how they manifest in their embodied forms in users, but initial results suggest that intellectual engagement largely manifests as focused attention, with relatively little physical activity.

We propose that it can be advantageous to enhance analysis of such segments of data by visualisation in three dimensions. Naturalistic, situated movement analysis will often include situations where there are long periods of stillness, as well as very small adjustments. By augmenting analysis methodologies, it is possible to extend the usefulness of motion capture technologies for measuring naturalistic, embodied affect.

5. CONCLUSIONS

Our 3D model allowed for concurrent comparison of the correlation plots of an entire dataset, which we have found useful for demonstrating the validity of our findings in this example, i.e. the excellent correlations between Vicon and raw accelerometer data, using linear transformation. Visual comparison of the three-dimensionally modelled first and second data sets (the genuinely correlated dataset and the negative control set, generated by reversing time series of the accelerometry signal) reveals clear differences between true correlations and correlations that are false, although they might appear statistically significant.

The next step will be to validate the model with other types of datasets. We are currently exploring the use of various plantar pressure data sets from the clinical biomechanics facility at Staffordshire University, where experiments have already been undertaken to investigate plantar pressure pattern during walking with bare feet and with various types of footwear. These experiments were simultaneously recorded with the Vicon opto-electronic camera system mentioned above. Visualised correlation of the plantar pressure data and the Vicon motion capture data may, as with the motion capture data we have used in the example we have used for this study, add an interpretative

dimension to the data and/or facilitate identification of true correlations.

6. REFERENCES

- Balaban, C.D., Cohn, J., Redfern, M.S., Prinkey, J., Stripling, R., Hoffer, M. (2004) Postural Control as a Probe for Cognitive State: Exploiting Human Information Processing to Enhance Performance. *International Journal of Human-Computer Interaction* 17, 275–286.
- Barker, S., Craik, R., Freedman, W., Herrmann, N., Hillstrom, H. (2006) Accuracy, reliability, and validity of a spatiotemporal gait analysis system. *Medical engineering & physics* 28, 460–467.
- Bland, J.M., Altman, D.G. (2010) Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies* 47, 931–936.
- Calvo, R.A., D'Mello, S. (2010) Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *Affective Computing, IEEE Transactions on* 1, 18–37.
- D'Mello, S., Dale, R., Graesser, A. (2011) Disequilibrium in the mind, disharmony in the body. *Cognition & Emotion* 26, 362–374.
- D'Mello, S.K., Graesser, A. (2010) Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction* 20, 147–187.
- Mota, S., Picard, R.W. (2003) Automated Posture Analysis for Detecting Learner's Interest Level, in: *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW '03. Conference On*. p. 49.
- Villanueva, M.B.G., Jonai, H., Sotoyama, M., Saito, S. (2000) Evaluation of the Ergonomic Aspects of Portable Personal Computers with Flat Panel Displays (PC-FPDs). *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 44, 654–657.
- Witchel, H.J., Westling, C., Healy, A., Chockalingam, N., Needham, R. (2012) Comparing four technologies for measuring postural micromovements during monitor engagement, in: *Proceedings of the 30th European Conference on Cognitive Ergonomics*. pp. 189–192.